

Learning How to Ask: Querying LMs with Mixtures of Soft Prompts

Year: 2021 | Citations: 577 | Authors: Guanghui Qin, J. Eisner

Abstract

Natural-language prompts have recently been used to coax pretrained language models into performing other AI tasks, using a fill-in-the-blank paradigm (Petroni et al., 2019) or a few-shot extrapolation paradigm (Brown et al., 2020). For example, language models retain factual knowledge from their training corpora that can be extracted by asking them to “fill in the blank” in a sentential prompt. However, where does this prompt come from? We explore the idea of learning prompts by gradient descent—either fine-tuning prompts taken from previous work, or starting from random initialization. Our prompts consist of “soft words,” i.e., continuous vectors that are not necessarily word type embeddings from the language model. Furthermore, for each task, we optimize a mixture of prompts, learning which prompts are most effective and how to ensemble them. Across multiple English LMs and tasks, our approach hugely outperforms previous methods, showing that the implicit factual knowledge in language models was previously underestimated. Moreover, this knowledge is cheap to elicit: random initialization is nearly as good as informed initialization.